AD-A231 541

A RAND NOTE

COPY

The Role of Uncertainty in Assessing the NATO/Pact Central Region Balance

Paul K. Davis

December 1988

DTIC
ELECTE
FEB 0 1 1991
S B D

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

40 Years 1948-1988

RAND

91

The research described in this report was supported by The RAND Corporation using its own research funds.

The RAND Publication Series: The Report is the principal publication documenting and transmitting RAND's major research findings and final research results. The RAND Note reports other outputs of sponsored research for general distribution. Publications of The RAND Corporation do not necessarily reflect the opinions or policies of the sponsors of RAND research.

Published by The RAND Corporation
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90406-2138

A RAND NOTE

N-2839-RC

The Role of Uncertainty in Assessing the NATO/Pact Central Region Balance

Paul K. Davis

December 1988

A research publication from
The RAND Strategy Assessment Center

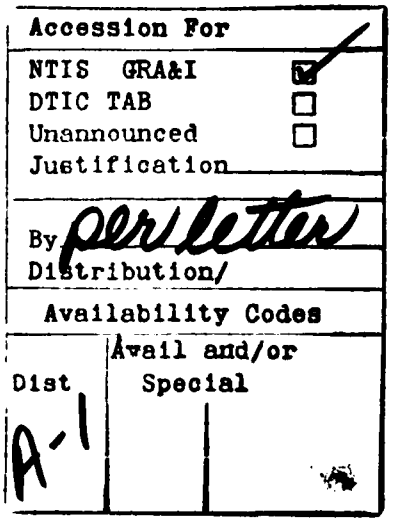
40 Years
1948-1988
RAND

PREFACE

This Note is a revised version of a paper developed originally for the Conventional Defense Study Group (CDSG) created by the Congress under the 1988-1989 National Defense Authorization Act. The CDSG is chaired by the Comptroller General and has representatives from the Congressional Budget Office, Congressional Research Service, Office of Technology Assessment, and General Accounting Office. The author was asked to develop, and present at a workshop, a paper on a net assessment of the NATO/Pact Central Region balance that would include issues of quality, readiness, mobilization, and sustainability. The original paper was published in April 1988 as RAND P-7427.

The changes incorporated here include editorial improvements, clarification of NATO's sustainability problems, and a better description of the analytic differences among competing approaches to balance assessment.

The research underlying the Note was conducted in the RAND Strategy Assessment Center under RAND's National Defense Research Institute, a Federally Funded Research and Development Center (FFRDC) sponsored by the Office of the Secretary of Defense. The Note itself, however, is the sole responsibility of the author and includes many subjective judgments.



SUMMARY

A basic question in any discussion of the military balance is which balance one is addressing: the balance of "inputs" such as defense expenditures and manpower under arms, the balance of combat equipment such as tanks, the balance of force readiness and mobilization potential, or the balance as measured by likely war outcomes if deterrence fails. This Note is concerned primarily with the last of these, and with the challenge of addressing that warfighting balance in the face of massive uncertainty rendering it meaningless to talk about allegedly "best-estimate" scenarios.

The beginning of wisdom about this balance is recognizing that war outcomes are sensitive to *scores* of factors, rather than the handful regularly discussed. Assessment should consider a vast range of plausible scenarios, where "scenario" is construed broadly to mean a set of assumptions about, for example, political-military context, warning times, mobilization times, alliances, operational strategies, force effectiveness, sheer quality of leaders and their troops, and even the "laws" of combat that determine rates of advance and attrition. Moreover, analysis should be based on a gaming approach, at least in structure, because the confrontation of opposing strategies and tactics is *fundamental* to warfare, and real wars seldom look like those in standard planning scenarios.

The results of such multiscenario analytic war gaming defy reductionist analysis: Simulated war outcomes often change drastically with what might naively be considered to be small changes in assumptions, and even the relative value of alternative improvement measures varies substantially from scenario to scenario. Measures or capabilities critical in some circumstances are almost irrelevant in others.

These wild fluctuations are not analytic artifacts, but rather a manifestation of something that professional military officers and historians have known since time immemorial, that war is an incredibly complex phenomenon characterized by uncertainty—except, for example, in instances where one side has overwhelming force (a situation that does not obtain in Europe). Moreover, tactics, strategy, and other human factors matter *greatly*. Caution should be exercised in using common analytic methodologies that obscure these basic aspects of warfare in the search for well-behaved and simply explained results.

What can be said from initial experience with multiscenario analytic war gaming applied to Europe's Central Region? The following are some personal conclusions:

negotiations. Arms control focused on readiness and other operational issues has the potential virtually to eliminate the threat of a short-mobilization attack (including variants in which the Soviets conduct extensive premobilization training). Such scenarios should probably be the most worrisome currently for NATO. Going further, force reductions could be stabilizing *if* sufficiently large and strongly asymmetric, but dangerous otherwise. The long-discussed move toward "defensive defenses" by both sides is also more worthy of study than it has been in the past.

- The long-standing policy of seeking conventional capabilities for merely an *initial* defense is now bankrupt: although a nuclear deterrent should be maintained, NATO's policy should include providing for successful conventional defense with no qualifications. This will require greater stockpiles to sustain combat for an initial and probably decisive phase and rapidly mobilizable production facilities appropriate for sustaining a longer war.

Turning again to methodology, I would argue that studies of the warfighting balance and arms-control alternatives should be based on both human war gaming and operationally sensitive simulation, because simpler treatments—however useful for communicating individual concepts—tend to omit many of the most important factors in actual warfare such as: maneuver phenomena; strategies employing surprise and deception; realistically imperfect decisions and behavior; and important aspects of readiness, mobilization, and sustainability. At the same time, even the more sophisticated games and simulations depend sensitively on uncertain assumptions about force-generation rates, the scoring of weapon systems, the scoring or nonscoring of support-force contributions to effectiveness, and other factors highlighted in recent years in both the classified and unclassified literature. If analysis is to serve the purposes of policy, these assumptions and others (many of which are NATO-favorable) need to be examined more critically than in years past. Also, the simulations themselves must be thoroughly understood.

Finally, Fig. S.1 provides a "fault tree" depiction of ways in which NATO could lose a Central Region war and suggests a systematic way of identifying and assessing measures to improve the military balance. It shows alternative ways for NATO's defense to fail, several of which have little to do with the quality of NATO forces in pure attrition warfare. The challenge, of course, is to block the paths to failure by eliminating deficiencies and other vulnerabilities.

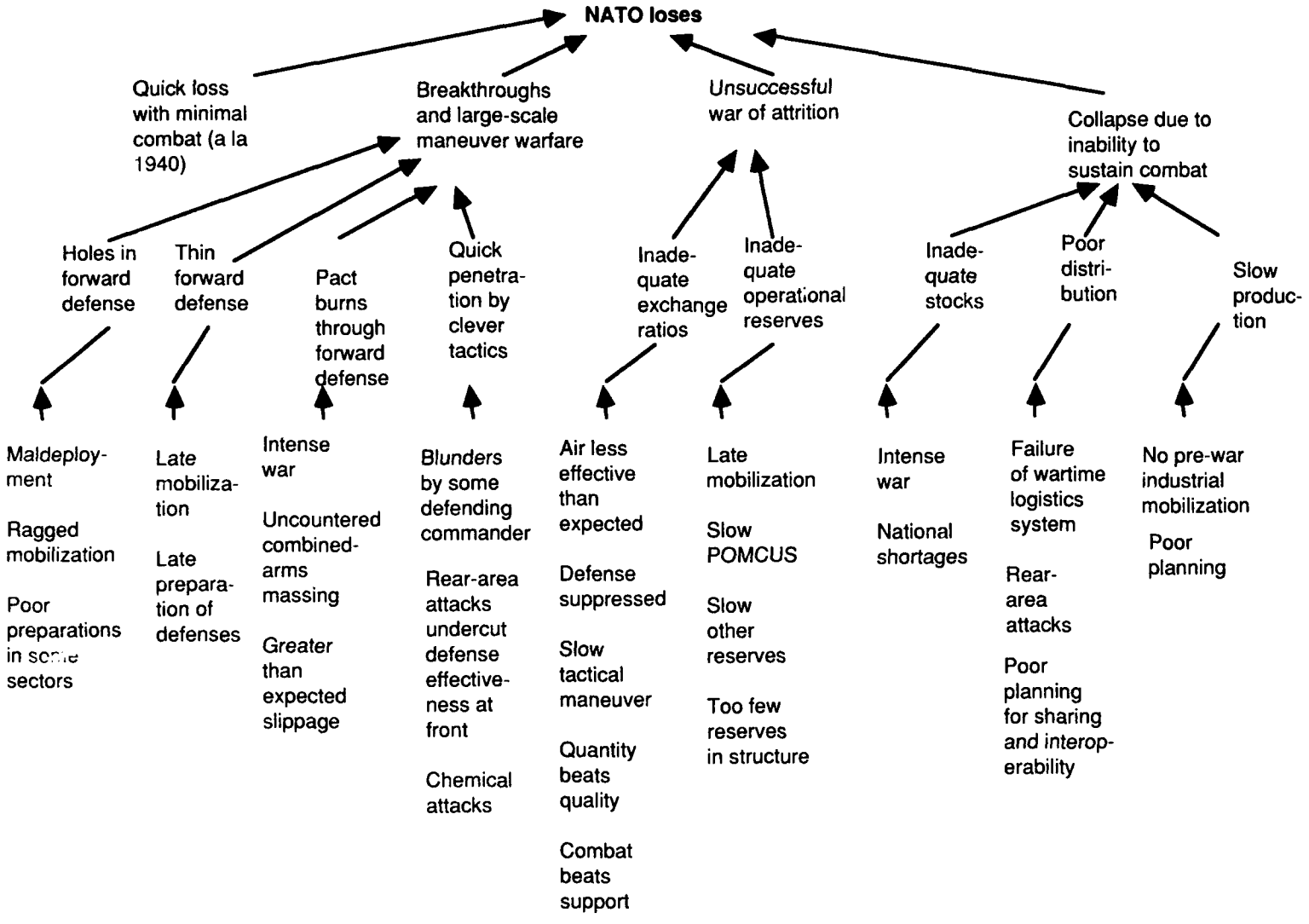


Fig. S.1— A fault-tree depiction of potential NATO problems

ACKNOWLEDGMENTS

The author has benefited greatly in his understanding of the Central Region balance from collaborative studies with colleague Robert Howe. Andrew Marshall has been a major impetus behind the effort to develop analytic war gaming as a methodology for improving balance assessments. Bruce Pirnie and Glenn Kent provided reviews with many thoughtful suggestions.

CONTENTS

FIGURES

TABLES

I. INTRODUCTION

DEFINING THE MILITARY BALANCE

Balance assessments are important.[1] They are important because they affect our intuition and our mindsets, which in turn affect indirectly everything from defense programs to operational planning. On one extreme, excessive pessimism about the balance can paralyze efforts to improve it, as suggested by the familiar refrain, "Why throw good money after bad? Everyone knows that conventional defense is not feasible and deterrence depends on the threat of nuclear retaliation." This deeply pessimistic image has been commonplace over the years and remains dominant within parts of the military community and in much of Western Europe. The image has been created primarily by widespread quotations about asymmetries in the "bean count" (e.g., the 2.4:1 ratio of main battle tanks in the Central Region).[2] It has been reinforced by General Rogers' statements, when he was Supreme Allied Commander for Europe (SACEUR), that in the event of war he would have to request nuclear-release authorization after a week or so, and by a continual stream of Department of Defense (DoD) briefings and studies in which the only question seems to be when (not whether) NATO's defense would crumble. It has been further exacerbated by the apparent failure, despite the Reagan Administration's defense buildup, to improve NATO's sustainability. "What do you *mean* we still can't fight for thirty days?" and "What have you done with all that money?" are common questions.[3] As I shall discuss later, my own conclusion is that the balance is much less adverse than the pessimists would have it.

At the other extreme, optimism or complacency can divert resources and postpone important problem solving. Overly enthusiastic assessments could also have a deleterious effect on NATO's planning for conventional arms control, which should be prudently conservative given the potential for the wrong type of agreements to reduce NATO's security.[4] In my view, some of the recent unclassified papers and books on the Central

---

[1]This section draws on material presented to the German Strategy Forum (Davis, 1985) and to a RAND conference on "Enhancing NATO Conventional Defense in Central Europe," March 3-5, 1986, held in Santa Monica, California.

[2]Levin (1988) provides a critically organized summary of static comparisons and explains how the comparisons can be made to appear much more or less adverse depending on details.

[3]Part of the answer is that the assumed rates of consumption have been greatly increased during the same period in which stocks were increasing (Shilling, 1988).

[4]This statement is based on unpublished 1986 work by the author, and on Thomson and Gantz (1987).

Region balance provide an outrageously rosy picture—however much I may agree with many of their arguments and deplore the tendency of others to exaggerate the threat. This is a complicated subject area where competent people of good will can and do disagree. One purpose of this Note is to explain why the disagreements exist.

One of the complications in the debate about the military balance is that there are actually many balances, some of them favorable and some of them unfavorable. As emphasized for many years by Steven Canby and others, NATO's *input* (spending levels and men under arms) compares well to the Pact's, but the bean-count *output* of combat systems strongly favors the Pact (see, for example, Levin, 1988, Karber, 1984, and Donnelly, 1983). However, NATO's readiness and support capabilities are in many respects superior (Levin, 1988, Posen, 1988), but the Pact's command and control is probably more cohesive. And so on.

This Note deals with a particular military balance, notably the warfighting balance that one infers by considering the likely and plausible outcomes of war if deterrence failed and the sides fought in a variety of scenarios with their *actual* forces, doctrine, command-control systems, and likely strategies. This balance does not *assume* a "fair fight," but rather assumes that the purpose of strategy, operational art, and tactics is to create especially favorable circumstances for battle.

TRADITIONAL POLICY-LEVEL ASSESSMENTS

Basic Buildup-Curve Methodologies. Before recommending a new methodology, it is useful to review what has been used in years past by those attempting to go beyond bean counts in performing analyses to be used in the development of broad defense policy and programs. The most influential of these, based on buildup curves, has had its locus in the Office of the Secretary of Defense (Program Analyses and Evaluation) (OSD/PA&E) and dates back to the 1960s. Over the years, many alumni of that organization, including myself, have published articles revealing the essential features of the approach. Other organizations such as the Congressional Budget Office (CBO) have also used it, and some individuals in the academic community have both used it and, in some instances, extended it.[5] The continuing themes of that school include (Davis, 1985):

[5]William Kaufmann, one of the nation's most experienced defense analysts and an adviser to numerous Secretaries of Defense, has also used, developed, and taught these methods for years (e.g., Kaufmann, 1983).

- The NATO balance is driven by assumptions about which nations will commit forces, which forces of each nation to include, how to count forces of different quality and composition, and timing (e.g., Blaker and Hamilton, 1977).

- NATO's chances for success should be reasonably good for theater force ratios less than about 1.5, with force ratios of 2.0 being quite worrisome.[6]

- The principal problem, then, is for NATO to assure that theater force ratios be kept as low as possible at all times—thus implying a need not only for substantial ready forces, but also for reserves forces adequate to support rapid mobilization and deployment.

- High-leverage measures include: (a) maintaining European reserves at a high state of readiness; (b) prepositioning equipment for U.S. forces so that fully equipped divisions can be available as quickly as the men can be flown in from the United States (POMCUS programs); (c) starting NATO mobilization early; and (d) obtaining substantial early French participation.

An additional theme emphasized for at least 15 years has been the argument that *either* NATO armies should reallocate their resources to increase the "tooth-to-tail" ratio, *or* they should be willing to give themselves credit in divisional scores for the benefits of support forces such as those able to repair tanks close to the battlefield, provide command and control to improve maneuver and fire support, and maintain the flow of munitions to the active battle areas. Recent unclassified articles by Posen illustrate the significance of this issue articulately (e.g., Posen, 1985, 1988). Unfortunately, it is not clear which part of the either-or statement is most appropriate. Some analysts, like Steven Canby, have long argued that NATO force structure has an unreasonably low tooth-to-tail ratio (e.g., Canby, 1986), especially for a short war, and that the extra tail doesn't help much.[7]

Yet another continuing theme has been the discrepancy between the way NATO assesses its own ability to quickly mobilize and deploy low-readiness units (poor) and the way it assesses Pact capabilities to do so (good). These matters are also discussed at length by Posen (1988). The DoD's *Soviet Military Power* (1987) confirms that many

---

[6]Roughly speaking, these rules of thumb relate to the famous 3:1 criterion of local concentration as follows: Imagine, say, 40 NATO and 60 Pact divisions scattered evenly among 8 corps sectors (5 and 7.5 divisions per sector, for an overall force ratio of 1.5). The Pact could take its excess 20 divisions and concentrate them on main axes. With, for example, 2 to 3 main axes, the Pact could achieve local corps-level force ratios of 2.8 to 3.5 if NATO failed to detect and react by counterconcentrating.

[7]Another complication here is that the attacker has advantages with respect to support: He knows where his main-thrust axes will be and where intensity will be highest. By contrast, the defender must have a logistics system adequate to shift both forces and munitions to where they are needed. I have not quantified this effect, but it is nontrivial.

Soviet and Pact units are currently at low states of readiness but provides few details. Levin (1988, p. 22) provides more details from an unofficial source of unspecified validity (Almquist, 1987).

None of the above topics needs to be examined with anything more sophisticated than a method for normalizing divisions to a standard measure (e.g., Armored Division Equivalents, ADE) and a model for predicting the rates at which various forces can be mobilized and deployed to the front. Such "models" can be back-of-the-envelope constructs plus some relatively detailed data tables distinguishing among units at different levels of readiness, although in the current era it is more convenient at a minimum to use a personal computer and spreadsheet software.

These simple models have been influential because they are understandable, dealing with issues at only the most aggregated of levels. Also, the principal conclusions drawn from them about improvement measures have been valid: Strategic mobility is good; rapid mobilization is good; operational reserves are good; and providing divisions in Europe with substantial firepower and mobility is good.[8] The models have been good enough to communicate a sense of how important these measures could be.

Extensions and Dubious Improvements. In recent years, there have been attempts to increase the sophistication of simplified analyses. Kugler, and more recently Posen, have, for example, used "FEBA-Expansion Models" that include force-to-space considerations and distinguish among NATO corps sectors. Kugler's original model was used heuristically to illustrate some of the difficulties in avoiding breakthroughs. Posen has used the model more aggressively in drawing broad conclusions about the balance (see Posen, 1985). There have also been a number of publications providing numerical calculations of attrition and movement based on solutions of Lanchester equations (e.g., Kaufmann, 1983, which discusses probabilities of successful defense) or improvements over Lanchester equations (Epstein, 1985).[9] For reasons discussed briefly in Appendix

[8]The simple model's emphasis on ADEs has been a chronic problem in some respects, however. In particular, zealous proponents of the firepower approach have often "proved" the nonutility of infantry divisions by showing they were costly per ADE procured. This was a spurious conclusion, because Europe is a complex theater with many types of terrain, and there are a number of specific but important zones where light but relatively mobile infantry could be far more effective than its firepower score would indicate, and in some ways more useful than armored forces. I encountered the analogous issue in earlier work on defense of Southwest Asia in 1979-1981.

[9]Another effect has been to encourage a cottage industry of individuals rediscovering the Lanchester equations and their intricacies. In some respects this is unfortunate, since the shortcomings of Lanchester equations are much more severe than is suggested by Epstein (1985)

A, I believe these intendedly more sophisticated aggregated attrition models have increasingly become more obscurational than helpful and do not represent a good direction for further research, although they may continue to be helpful for communicating certain key ideas.[10] The basic problem is that adding content to reflect correctly maneuver issues, command-control problems, breakpoints, flank protection, and other matters quickly becomes complex and data-intensive unless one does violence to the underlying phenomena. Despite my general preference for simple analytic models, I would argue that the next natural step beyond analyses focused on forces and force ratios vs. time is careful simulation and war gaming (with human involvement). These are not, however, efforts to be undertaken lightly, and classified information and the participation of military experts is often important. Also, the efforts are counterproductive unless the models remain understandable.

SHORTCOMINGS AND CHALLENGES

Before leaving this description of traditional policy-level analytic methodologies, with and without extensions, I might emphasize that one of the school's historically rooted ethics has been that policymakers should assure that the military has adequate resources and, when necessary, "prove" that it has them (e.g., as in Enthoven and Smith, 1971), but not explore how the resources (forces) should be employed or what problems might be encountered in doing so. Instead, such operational analysis should be, in this ethic, the responsibility of the generals. This sounds virtuous, but it ignores the strong relationships among politics, policy, strategy, doctrine, training, exercises, and warfighting. There are many political constraints on what NATO's generals are permitted to plan and exercise in peacetime, and many of them severely undercut military effectiveness.[11]

Despite the many advantages and the track record of success of such methods in the past,[12] these superaggregated analyses present several problems. Among the more serious are:

---

or more recent articles (e.g., Lepingwell, 1987, and Homer-Dixon, 1987), and such investments of effort in Lanchester theories are probably misplaced (see Appendix A).

[10] I note that Mearsheimer (1988, footnote 5) is also dubious about the extensions of methodology and that Kugler has based much of his recent work on human war gaming.

[11] One of the more vociferous, hyperbolic, but thoughtful commentators on such matters has been Steven Canby. See, for example, Canby, 1986, and references therein.

[12] Analyses based largely on theater force ratio vs. time have been influential in decisions to buy POMCUS equipment, strategic mobility assets oriented toward the Rapid Deployment

- They lack credibility among those familiar with more detailed treatments, who recognize that many of the key issues are misrepresented or overlooked.

- Because of their aggregation and associated abstraction, they have no potential for unifying such disparate communities as the technologists, historians, maneuver warfare advocates, and resource managers. They are read by one community and ignored by the others.

- They have little useful to say about matters of operational strategy, command-control, doctrine, logistics, sustainability, force composition, the potential value of new weapon systems, or (with some important exceptions) the relative merits of alternative arms control measures.

The last item is especially damning, since it is most unlikely that NATO will be in the force *expansion* business in the next decade. Instead, in the absence of favorable forms of arms control on a large scale, **most improvements in the balance will come about precisely because of improvements in the very things that the buildup-curve analysis is poor at capturing. So also are arms-control studies requiring the more detailed analysis.**

There is also a tendency in using the methodologies to confuse what the balance "should be" considering NATO's inputs (and even its output resources) and what the balance actually is—as measured by what would happen if war occurred. Even if there are references to likely war outcome that suggest interest in the warfighting balance, the tendency is to assume away the effects of surprise, intra-alliance coordination problems, doctrinal and other constraints, the inherent advantages of the aggressor, and likely defender mistakes. These issues may be acknowledged in footnotes, but usually in terms of noting "different problems" having nothing to do with "the balance" as the authors conceive it.[13] Again, then, disagreements about the balance often begin with different conceptions of what "the balance" means.

Force, and prepositioned equipment for Southwest Asia (Davis, 1982). They have also been influential in discussions of the military balance in Europe, Korea, Southwest Asia, and elsewhere.

[13]See, for example, Mearsheimer (1988, footnote 10), who is at least explicit on this matter.

II. MULTISCENARIO ANALYTIC WAR GAMING

BASIC PRINCIPLES

As discussed elsewhere (e.g., Davis, 1985, 1987), the approach to balance studies and other matters being taken by my colleagues and myself in the RAND Strategy Assessment Center (RSAC) is an attempt to meld the better features of human war gaming and analytic modeling. The war-game style is especially important for bringing in a wider range of variables and complications, and for assuring that analysis confronts issues of strategy and tactics under conditions of imperfect information. The work depends heavily on combat simulation models[1] and decision models (see Appendix B). Most of the work is interactive, with military analysts playing through simulated wars in some detail. In one common mode of operations an analyst may play Blue, entering orders in an attempt to defeat an automated Red commander following a plausible Red strategy with Red doctrine and forces.

For the purposes of this paper the most important aspect of the RSAC work is our emphasis on facing up to massive uncertainty—i.e., on highlighting the scenario variable. Consider, for a moment, a baseline case in standard analysis, a case that is often treated as though it were a best-estimate scenario: (a) on Pact M-day, the Soviet Union and all its Warsaw Pact allies mobilize and prepare for war with all of the non-Soviet Warsaw Pact states charging forward enthusiastically with their Soviet masters (probably a rather pessimistic assumption for NATO); (b) on NATO M-day, the United States and all its NATO allies mobilize together and proceed without friction to implement their war plans (probably an optimistic assumption for an alliance of independent nations reacting to ambiguous warning), (c) war occurs without surprises (e.g., weapons work as advertised, strategies are as advertised or anticipated, and attrition warfare prevails with an elastic defense line), and so on. Whatever this scenario represents, it is not a best estimate, but rather some bizarre mixture of various optimistic and pessimistic (and often unrealistic) assumptions.

The answer, it might seem, would be to construct a realistic planning scenario—a true best-estimate scenario. If one thinks about how to do this for awhile, the difficulties

[1]There is a long history on the development of theater-level combat simulation models. Some of the better known over the years have been ATLAS, CEM, LULEJIAN, COMBAT II, IDAGAM, TACWAR, VECTOR, IDAHEX, TOTEM, MASTER, and RSAS (and its combat model component CAMPAIGN). There is no single reference on these, which vary greatly in many respects, but some of the earlier models are described in Battilega and Grange (eds.), 1984.

become clear—at least to most people. (There are also those who cling to the notion of a best-estimate scenario even though the quality of the best estimate would be very low.)

The alternative is to explore a diversity of scenarios in an attempt fully to face up to uncertainty. Table 1 illustrates the questions one may ask.

Table 1

ILLUSTRATIVE "WHAT IF?" QUESTIONS

- What if one or more of the NATO allies reacted slowly in crisis, resulting in a ragged mobilization process and disrupting the general defense plan?
- What if Poland cooperated only minimally with the Soviet Union or fought with less than high intensity?
- What if Pact forces proved somewhat less effective, for constant equipment, than was nominally assumed?
- What if the intensity of war proved higher (or lower) than is usually assumed? Would NATO fare better or worse, and how would this affect sustainability?[2]
- What if deployment times proved much longer than is usually assumed (e.g., for U.S. POMCUS forces or low-readiness Soviet divisions)?
- What would a Soviet simulation of conflict look like if it began by assuming a NATO invasion of the Pact? How would this affect the circumstances of battle and the nature of campaigns?

DIMENSIONS OF UNCERTAINTY

If Table 1 is enough to put the reader in the right spirit for multiscenario analysis, then Table 2 will be understood as an effort to be more systematic. It summarizes the key dimensions of uncertainty that my colleagues and I try to consider and includes an illustrative set of the specific variables that can be treated in studies of the Central Region. **In our parlance, a "scenario" (in the context of discussing "multiscenario analysis") is a set of assumptions about all of the various issues treated in the table**—issues ranging from the political-military scenario to the value of certain technical parameters or even the equations that should be used in the simulation models of warfare.

It should be evident from even a brief perusal of Table 2 that there is an enormous number of possible cases. In practice, we must select cases that seem likely to be fruitful

[2]Loss rates, which is what I mean here by intensity, are correlated with but different from consumption rates.

Analysis, then, becomes something of an art—with the potential for providing great insights or causing great mischief. It is no panacea.

Table 2
DIMENSIONS OF MULTISCENARIO ANALYSIS
(with illustrative variables)

Political-Military Scenario	Strategy and Tactics	Force Structure	Technical Factors
Number of theaters	Duration of Soviet mobilization	Size of threat to Central Region	Intensity of war i.e., attrition rates)
Time between wars in different theaters	Soviet scheme of maneuver and deception	Extra divisions Fewer divisions	Densities at which breakthroughs occur
NATO mobilization times by theater	NATO defense strategy by circumstance	Arms-control reductions to both sides	Tacair effectiveness for killing and countermaneuver
Allied behaviors (both alliances)	Soviet use of other theater	Readiness levels	Helicopter effectiveness
Premobilization preparations	Use of air forces	Days of supply	National fighting effectiveness
			Unit breakpoints by readiness level
			Value of support forces (e.g., repair and C^3I)

SENSITIVITY OF OUTCOMES

Naively, one might hope that many of the cases would prove uninteresting—that the sensitivity of simulated war outcomes to most variables would prove to be low. Unfortunately, that is not the case. To the contrary, simulated war outcomes can be highly sensitive to almost any of the variables in Table 2. As an example, simulated outcomes can flip from a victory for the Pact to a stalwart defense for NATO if one merely changes assumptions about the intensity of war—without ever leaving the range

of highly plausible loss rates (i.e., attrition rates), which is at least a factor of two and probably more like a factor of three.[3]

An even more troublesome reality is that even the *relative* value of different improvement measures, or the relative importance of different variables, depends strongly on the scenario (as defined by the value of the "other" variables). This is hardly surprising to someone who thinks for a moment, but it is disquieting to the analyst hoping to proceed by mechanically churning out excursions from a well-oiled model. As an example here, consider the value of a postulated new weapon system and concept of operations for interdicting the Pact's deploying forces. In short-mobilization scenarios in which the sides are both scrambling in a mobilization and deployment race, one might expect the payoff for interdiction to be high, but in scenarios in which there have been weeks or months of preparations and movement, it might be modest. This situation-dependence of sensitivities is the rule rather than the exception, and analysis is difficult and technically complex.

To provide some insight about the outcomes of such analytic war gaming, let me describe qualitatively part of a typical presentation. First, we show results for one of the several baseline scenarios. Suppose that this happens to be one of the scenarios in which NATO does well—holding at or close to the border, extracting a highly favorable exchange ratio, and not running out of supplies for the duration of the war simulated. We describe in some detail why the scenario is plausible, what assumptions resulted in the particular outcome, and why the result is not as outlandish as some might think initially (because they have usually been briefed only on unfavorable cases). Next, however, we drop "the other shoe." We begin introducing highly plausible assumptions that change the scenario "slightly." Perhaps we assume that one or another of the NATO allies mobilizes more slowly than the others, or that U.S. POMCUS forces deploy slowly, or that the Soviets are aware of some special weaknesses in a particular area, or that the Soviets have a rather successful H-Hour air strike. Or perhaps we change the underlying equations of the simulation "a bit," and perhaps some of the parameters in those equations. Suddenly we find that what was previously a "good scenario" collapses: NATO loses the war quickly and decisively, as illustrated by Fig. 1. So much for optimism.

---

[3]This is based on the author's unpublished work using a variety of sources for information on World War II and the Arab-Israeli wars, including recently published Soviet attrition data from the Eastern Front campaign (Stoeckli, 1985). See also Posen (1985),

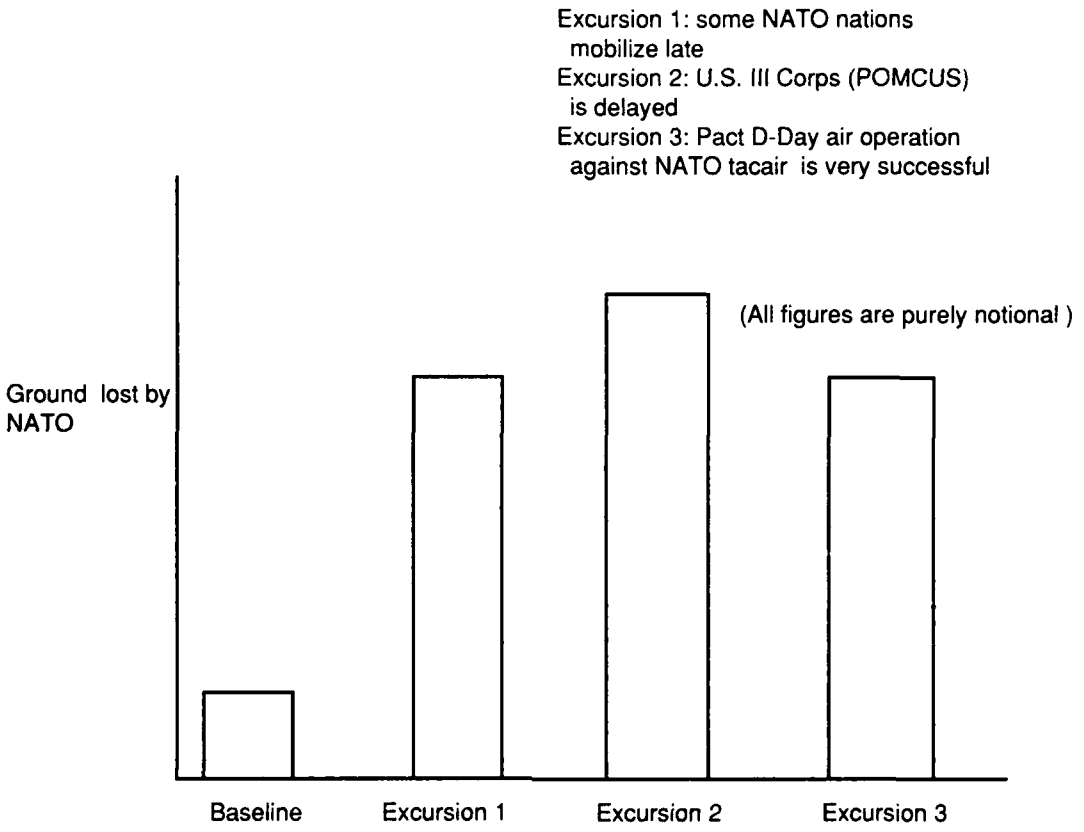


Fig. 1—A defense-favorable case and excursions

Next, however, we look at possible Soviet planning. A second baseline case might be one that appears well suited to Soviet objectives and attitudes about warfare. In the baseline case the Pact may do extremely well, with a campaign that looks like a textbook case of Soviet doctrine—early breakthroughs, vigorous exploitation, and so on. It is not difficult to construct such cases. Again, however, we drop the other shoe. We imagine ourselves to be a prudent Soviet planner and consider some of the "What ifs?" that might trouble him. For example, NATO might vary its operational strategy somewhat—compensating for the maldeployment that is well known to students of the balance (e.g., Levin, 1988, p. 8). Or it might be that non-Soviet Warsaw-Pact forces fight with somewhat less enthusiasm and effectiveness than their equipment might suggest. Or, it might be that the Soviets make more pessimistic assumptions about the quality of their own equipment as reflected in divisional scores (especially for older equipment). And so on. Again, we find that the results of the baseline case collapse, and what was originally a very favorable scenario for the Pact turns into an unmitigated disaster (Fig. 2).

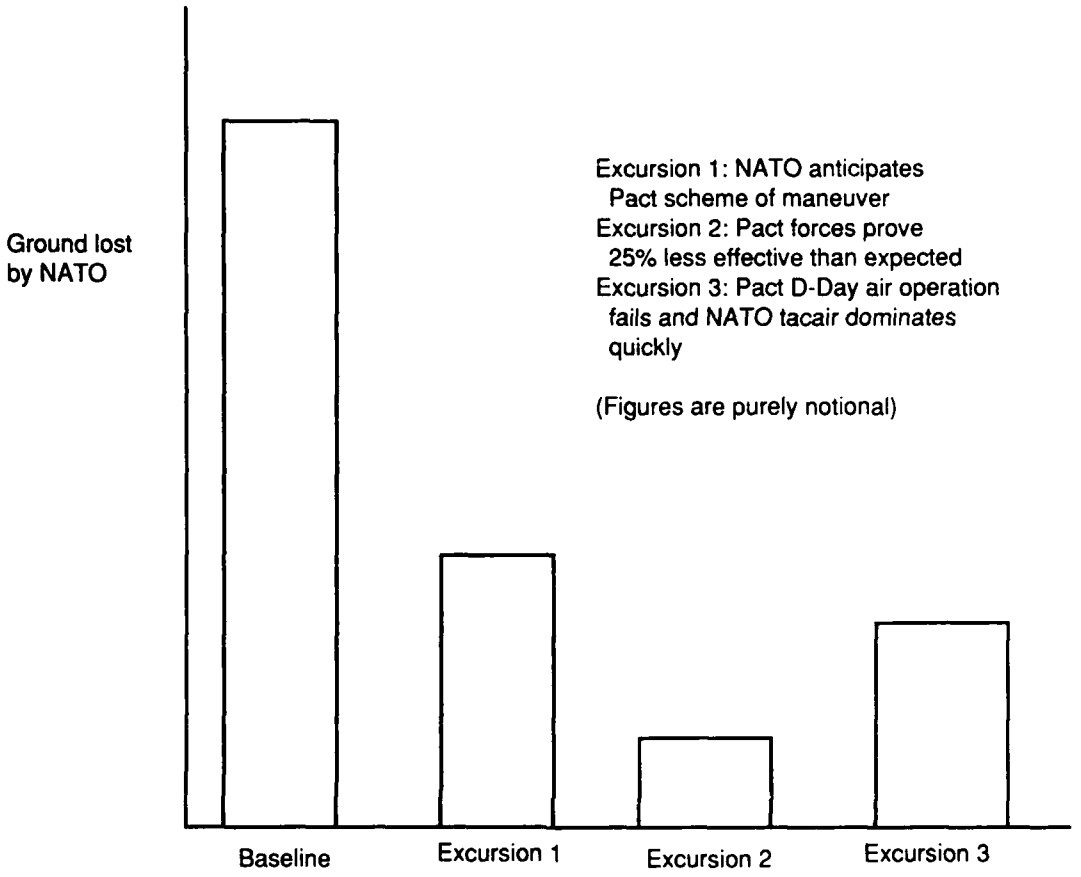


Fig. 2—An illustrative offense-favorable family of excursions

In practice, many of our presentations exploit map-based depictions of the campaigns, and these are more dramatic in showing how success can flip into failure. Upon occasion, we can even find instances in which a NATO counteroffensive is plausible.

It would be pleasant to report that results come out with NATO doing well at least half the time. Such is not the case, although at this point I don't know what "half the time" means. because neither I nor anyone else knows how to weight the probability of the various cases. NATO suffers from a significantly adverse force ratio[4] and all the disadvantages of granting initiative to the Warsaw Pact, which can pick the main thrusts of attack, mass fire, and orchestrate the timing of at least initial operations. Also, many aspects of combat modeling tend to introduce defense-optimistic biases. Or, to put it

[4]In reviewing old studies, I noted that in the early 1970s a theater force ratio of 1.5 :1 was considered dangerous. From reading the current optimist papers, one might believe that the danger ratio is now 2:1, although the laws of war have not changed.

differently, my colleagues and I believe that many standard assumptions favor the defense implicitly and are valid only if the defense is ready, competent, and relatively unconstrained—something that should perhaps not be assumed for a complicated coalitional force in the first days or weeks of a war decades after the last comparable war that breaks out.

When we examine sensitivities dealing with these matters, NATO often suffers in the result because, simply, NATO lacks adequate operational reserves to compensate for things going wrong. The development of III Corps is a great help in this regard, at least for scenarios in which III Corps is able to deploy, but even so, NATO's conventional defense is anything but robust. In thinking about this, the reader should understand that **in the most optimistic assessments, it is assumed that NATO maintains a coherent defense at all times—something that should simply not be assumed, even if the line ought to hold "on average" according to commonly used analytic theories. Holding "on average" is not good enough if penetrations in any one area can be exploited to change the nature of warfare from defense-favorable attrition warfare to the large-scale maneuver warfare for which Soviet armies have been organized and trained.**[5]

ORDER OUT OF CHAOS

Multiscenario analysis can be confusing and even paralyzing. It would seem possible to "prove" anything by merely choosing among the set of possible reasonable assumptions. What, then, does one do? **Figure 3 is an example of a generic technique that I like to use in making sense out of multiscenario results. It is a "fault tree" depicting various paths to disaster. The challenge is to make the paths unlikely.** The idea, after all, is not to sit and wring one's hands about the fragility of the balance, but to find ways to improve one's odds. By looking at possible ways to fail, one can itemize issues for priority attention. In some cases, this consists of doing research to find out more accurately how well a particular system might work or how high loss rates would probably be. In other cases, it means buying things, changing operational plans, or

[5]Concerns about such matters have long been expressed by history-reading maneuver enthusiasts. One way I have used for some years to test related sensitivity in analysis is to *specify* a localized breakthrough in one or another sector at various times, and then observe how well the defense is able to react and contain it. I regard this as a very useful measure of effectiveness for the *robustness* of NATO's defense in various scenarios. Another technique is simply to specify that initial Pact movement rates are fast enough to get through the zone of prepared defenses quickly and to then simulate the consequences.

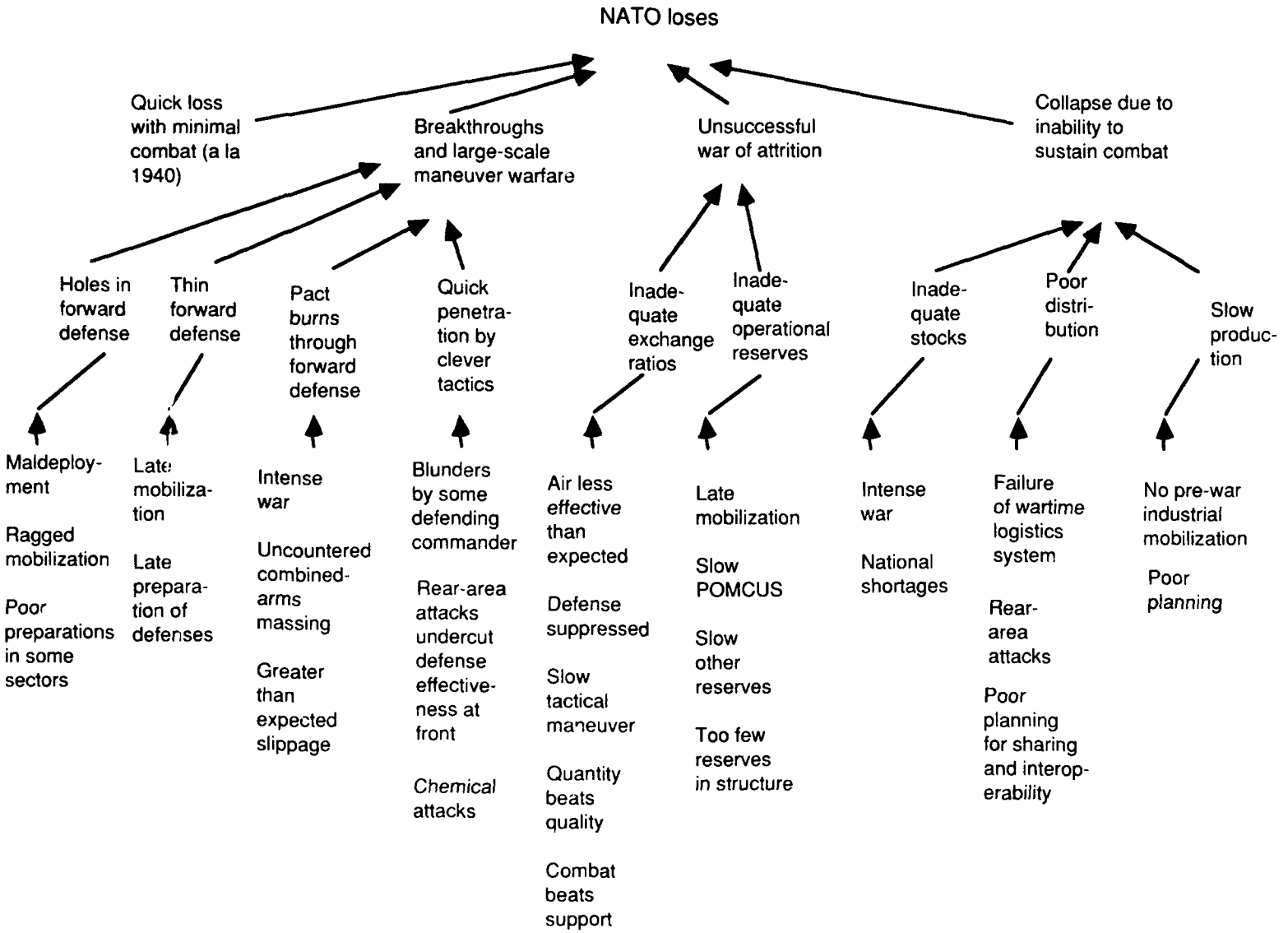


Fig. 3 —A fault-tree depiction of potential NATO problems

exercising the capability to improve the competence component of readiness. In still other cases, it could mean seeking arms-control provisions.

An important part of the method is to construct similar trees for the Soviet side so that we can better understand what strategies it might construct to improve its odds and what adverse circumstances it might be most concerned about. Such trees are not simple inversions of the NATO tree (except in the limit in which one makes both trees comprehensive—and incomprehensible), because the Soviet side has a different perspective and different variables over which it has control. Without going into details here, it should be noted that, as the aggressor, the Soviet planner has important advantages in his ability to shape and control the scenario—at least for the early period of conflict, and at least if he is able and willing to back off and try again later if initial efforts to shape an appropriate scenario fail because of NATO reactions.

To illustrate one line of reasoning, suppose that the Soviets wish to avoid a war of attrition because of the advantage defense has under such circumstances, given reasonable time for preparations.[6] This almost implies that the Soviets would seek to create a scenario in which either NATO as a whole has not mobilized and prepared its defenses in depth, or at least one corps sector has not been prepared and well covered by D day. Either scenario would virtually require achieving strategic surprise. At the same time, according to the DoD's *Soviet Military Power* (1987), Levin (1988), and other sources cited by Posen (1988), much of the Soviet army is currently at a low level of readiness. It might prove extremely unreliable under combat conditions if it does not receive *extensive* training (i.e., two months and, quite possibly, much more). But without such forces, the Soviets would lack the benefit of reserves to compensate for things going wrong, such as a bogging down of the first echelon. How, then, to proceed?

The answer seems clear to me, if not easy to achieve against a vigilant NATO. In my view, **the ideal scenario for a Soviet planner would be to prepare at least a significant portion of his forces over a long period of time**—gradually raising the readiness levels and reliability of enough "low-quality" units to assure adequate reserves. After such a period, which might be many months in duration after being triggered by a

[6]In this connection, remember that the Soviets all know and remember the battle of Kursk in World War II, a gargantuan battle in which they were on the *defense*. The German Army attacked and fought ferociously, but in spite of paying with high loss rates was unable to penetrate the prepared defenses in depth. The Soviets then launched a counteroffensive of historical significance. See, for example, Von Mellenthin (1955) or Caidin (1974). To the Soviets, attacking a fully mobilized NATO with forward defense in depth must surely resurrect the wrong type of image for optimism.

fundamental change in East-West relations, the Soviet planner would wish to orchestrate a short-mobilization attack that would be under way before NATO forces were all in place and that *might* even begin without a decisive and coherent NATO response to warning because of ambiguities in the situation and Soviet-generated hints that war might still be averted in the absence of precipitous and provocative actions such as a full-scale NATO mobilization.[7]

Every analyst has his own favorite threat scenario, but this is mine. It strikes me as the most plausible of the bunch, and dangerous indeed.[8] It bears little relationship to those discussed in the more optimistic balance assessments.

DISCUSSION

Using multiscenario analytic war gaming to identify improvement measures tends to bring out or dramatize many issues that would not even be treated in a more standard form of analysis. Instead of dwelling on well-known problems and the kinds of solutions that are visible in standard planning scenarios (e.g., "It would be nice if NATO had ten additional divisions"), it tends to identify a series of discrete, important, and solvable problems—many of them approachable with existing resources. Some of these problems have not previously been solved because the solutions have been deemed politically unpalatable, whether rightly or wrongly. One might hope that if the consequences of certain problems originating at the political level can be demonstrated, pressures will build to correct them—despite the need for negotiations and coordination. It is heartening to observe, for example, that the Levin report (Levin, 1988) correctly highlights the maldeployment of NATO forces as one of NATO's more serious problems.

In summary, then, this operationally sensitive multiscenario analysis is a fundamentally different way of approaching balance assessments and strategic planning. It faces up to uncertainty and emphasizes the importance of hedging,

---

[7]By no means do I wish to imply that this strategy would succeed. If NATO were reasonably vigilant, there would be many things it could do during the period of extended tension in response to the increases in Soviet readiness. It is unclear, of course, whether it would in fact do those things if the signals were ambiguous or the political leaders were distracted by other events. Nonetheless, there is much that *could* be done: adjustment of operations plans, creation of obstacles, higher states of day-to-day readiness, and even creation of additional operational reserves that would be available early.

[8]An excellent reference on surprise attacks is Betts (1982). Vigor (1983) is also quite interesting in its discussion of Soviet emphasis on surprise, although I do not find Vigor's favorite surprise-attack scenario convincing. Levin (1988) also discusses a short-warning scenario without prior preparations. Private communications with Christopher Donnelly suggest that he considers it more plausible, as I do, that a short-mobilization attack would come after extensive preparations.

adaptation, flexibility, aggressiveness, and other characteristics military commanders understand but have often not been able to explain or translate into specific action measures supported by political authorities.

Once one truly accepts and internalizes the paradigm, balance assessments can never be the same again: There is never a single answer, but only an abstract imprecisely defined concept such as that depicted in Fig. 4—a concept in which one knows that the breaks might fall one way or the other way, and the purpose of planning is to improve the odds.

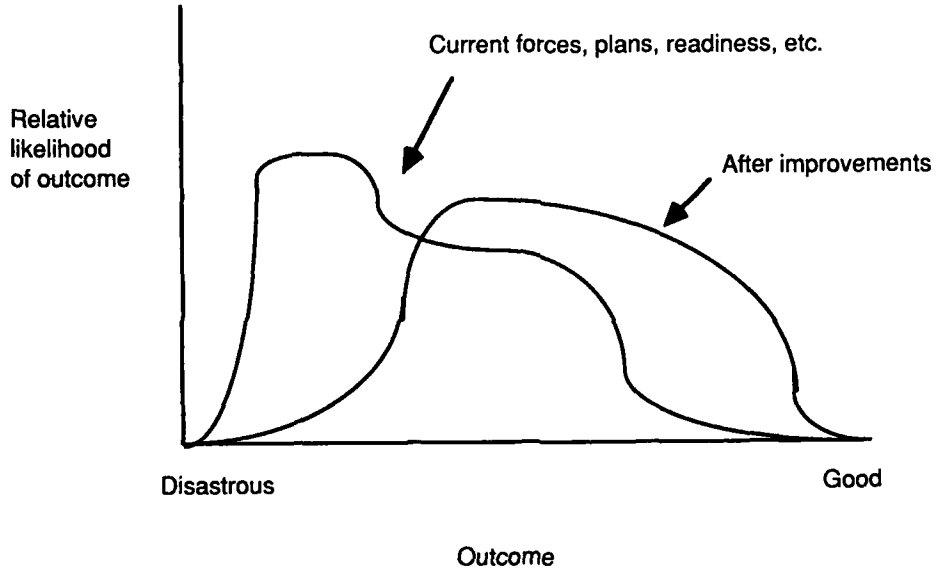


Fig. 4—A notional multiscenario image of how one thinks about the value of improvement measures

III. READINESS, MOBILIZATION, AND SUSTAINABILITY

Let us now turn to some specific issues of readiness, mobilization, and sustainability—although I have touched on many of them already.

READINESS

The word "readiness" is often construed rather narrowly and technically in military analysis, but in this Note I take the view that there are many aspects to readiness, some of them qualitative. Perhaps the most important, although certainly among the more fuzzy, is the issue of man-for-man fighting effectiveness for constant equipment. As we have seen from millenniums of history, wars can be won by the more proficient and innovative side despite quantitative inferiority. Generalship matters; doctrine matters; training matters; and so on.[1] These are not minor issues when it comes to predicting the outcome of wars. Indeed, it is simply not possible to understand the results of historical conflicts without facing up to such matters (Dupuy, 1979). For example, it has been estimated that German forces were twice as effective as their Russian counterparts on a man-for-man tank-for-tank basis in World War II at the tactical level (the Soviets performed very well at the operational-strategic level after the initial debacle). Similarly, only the most stubborn of technologists would deny the evidence that Israeli military forces are simply better, for equivalent equipment, than their Arab antagonists.

Are there examples of such imbalances in the modern NATO/Pact Central Region standoff? Undoubtedly there are, but it is difficult to predict most of them with any confidence. By and large, for example, most observers believe that U.S. air forces are greatly superior to Pact air forces in their competence for air-to-air operations. This comes from a combination of tradition, major investment in realistic training operations (and the willingness to accept losses), first-rate aircraft, and equally first-rate technology for command and control (as Israeli results in the Bekaa Valley indicated, such technology can pay off). Unfortunately, we are unable to assess with any degree of confidence how NATO's superiority in tactical air would translate into operational effectiveness on the ground. Also, the conditions of many-on-many combat in the Central Region are quite different from those in the Bekaa Valley.

---

[1]This point can be overdone. We should remember that the stronger side *eventually* won World War I, World War II, the U.S. Civil War, and the Napoleonic wars. "Strength" in such cases, however, was measured by total national capability to mobilize and conduct long wars, and not by capability to prevail in the first military campaign.

If we turn to ground forces, there is little that can currently be said with confidence beyond what appeared in Senator Levin's 1988 report. Training levels and morale in NATO forces appear higher than in Pact forces. But combat effectiveness depends on a host of factors, including doctrine, circumstances, and who has the initiative. It is notable, for example, that Arab forces fought with bravery and skill during the Yom Kippur war and did well in certain battles where they possessed the initiative.[2] Also, Soviet doctrine is designed to compensate for what the Soviets have long recognized to be a reality of their society—a lack of creativeness and initiative at lower levels of command. As long as the Pact had the initiative, the fighting effectiveness of Soviet troops might be high. Even in human war games, with all their uncertainties, one can see the tangible benefits to the Pact of having the initiative: The NATO commander is constantly trying to assess the situation and react, and is often "behind the power curve"; by contrast, the Pact commander can pursue his plan straightforwardly—until and unless things go badly awry.[3]

Turning from the philosophical to the specific, **how can we reflect readiness issues in theater-level analyses?**[4] With "readiness" used as it is here, to embrace many qualitative aspects of "effectiveness" that depends on practiced skills, the answer is that we reflect readiness issues implicitly or explicitly in many ways whenever we use games and simulations—e.g., in assumptions about:[5]

- Fighting effectiveness for constant equipment, as discussed above (to assume nothing is to assume something—that all forces are equally capable!); we currently assume that effectiveness scores increase linearly with training time from M-Day levels to "full-readiness" levels.

- Sortie rates, kill rates, and loss rates for fixed-wing aircraft and helicopters as a function of mission, type of battle, and time of day.

---

[2]Dupuy specifically identifies "setpiece battles" as special because brave but otherwise only modestly competent forces may be much more competent than in a free flowing situation.

[3]The same type of phenomenon occurs at the tactical level, and those experienced with the National Training Center can testify how dramatically the competence of defenders (and attackers) changes with experience—in this case synthetic experience.

[4]Not discussed here is the very important issue of top-level command-control readiness, which depends on such matters as the realism of theater-level exercises and the appropriateness of political and military decisionmaking arrangements for the circumstances of crisis and war. To explore such issues we consider, for example, scenarios in which NATO obtains strategic warning but acts upon it slowly and with "halfway measures."

[5]It may not be apparent that all of these depend on "readiness," but in fact they are all dependent on the broad version of readiness that includes qualitative effectiveness honed by preparation, practice (e.g., war gaming), and full-alert conditions.

- Delay times in theater- and corps-level command-control processes (e.g., how long does it take the theater commander to correctly identify the main thrusts and react? How long does it take him to recognize a potential breakthrough situation and allocate theater reserves to a corps in trouble? How long does it take him to reallocate and reapportion tactical aircraft?).

- What frontage can a division defend and still hold ground? (One might assume that in the absence of combat experience a division would initially find itself losing ground under circumstances where doctrine said it could hold—unless the attacker were equally green.)

- How far could a defensive division be stretched without suffering severe penetrations, local envelopments, and a breakthrough? (This would be a function of skill in reconnaissance and maneuver, familiarity with terrain, and ability to undercut enemy operations.)

- What operational strategies are plausible, given the distribution of readiness across units and the associated nominal and conservative buildup rates?

These examples illustrate how readiness enters into simulation assumptions and makes the point that **there are outlets for measuring the value of readiness if we make the effort to do so and accept the need to make subjective assessments.** In recent years, the intelligence community has been increasingly helpful on such matters. Nonetheless, there is much more that could be done, and many of the current outlets are not used or are used with dubious assumptions—for both Pact and NATO forces.

MOBILIZATION

There are many dimensions to mobilization, including the process of simply filling out units with warm bodies and the process of preparing those bodies to work effectively within their unit and in cooperation with other units (training). Training, of course, directly affects the readiness discussed above. The issues are not identical, however, because one might argue that even a rather lengthy mobilization would at best bring forces up to a nominal level of readiness consistent with the nation's doctrine and planning factors. One side's ready forces might still be much less effective than the other's ready forces.

As a minimum, however, we should assume that neither side can perform miracles and that it takes substantial time to turn civilians (even with prior military experience) into fighting men. It has been a continuing source of irritation to those of us attempting net assessments and defense planning over the years that the Soviets are generally given credit for being able to mobilize forces and send them into battle far more quickly than the NATO countries. Many people believe we give the Soviets too much credit in this

regard, even with the DoD reassessments alluded to in Posen (1988). I am unaware of any effort to perform a careful assessment using the same experts and measures to look at both sides' plans and capabilities.

One factor that has confused analysis in the past has been the Soviet doctrinal willingness to use lower-readiness divisions when necessary, even if they would be expected to take high losses. This is hardly surprising, considering the Soviets' history in the Great Patriotic War was that only by throwing everyone into the breach were they able, barely, to fend off defeat. Also, it is not surprising for a nation with a doctrine calling for (and postulating) early breakthroughs by first-echelon forces. If the Pact forces were indeed able to achieve early breakthroughs, it is plausible that low-quality forces could exploit those breakthroughs. Let us now consider, however, the case in which the Pact does not have a cakewalk—the case in which the first echelon is stopped cold by NATO's defenses, and the second echelon is asked to assault prepared defenses in depth, albeit defenses manned by battered NATO forces of uncertain cohesion and capability. Under these circumstances, I would expect low-readiness assault divisions to be ineffective—especially since NATO doctrine would be encouraging innovative and aggressive tactical (and conceivably operational-level) operations to frustrate Pact operations. Our simulations can reflect such effects straightforwardly.

If this chain of reasoning is valid, it seems likely that analysis of the balance should: (a) assume that Pact planners would attempt to bring second-echelon forces to a considerable level of readiness before attempting an invasion, even if this required months, and (b) assume that forces asked to assault prepared defenses in depth would be relatively ineffective in terms of man-for-man fighting capability and breakpoints unless trained for considerably longer than is often considered adequate to bring forces to full readiness. Similar assumptions should be made about the lower-readiness NATO units, but it is at least plausible that such units would be better in the defense than equally competent Pact forces would be on the offense. There is certainly historical basis for such a belief.

If, as I have argued elsewhere and Posen and others have argued in recent articles, the Soviets' second-echelon forces are at relatively low states of readiness that would require considerable time to change, then **an important objective of conventional arms control should be to place restrictions on Soviet efforts to increase readiness (Davis, forthcoming).** Although verification problems would surely be complex, it is important to recognize that significant increases in readiness would involve activities by many tens of thousands of people drawn from the civilian sector. One might expect that a Soviet

leader contemplating such activities would consider it likely that they would be observed. In an arms-control regime that included various types of intrusive inspection, the risks of covert training might be increased even more.

Another important point about mobilization is that it is not a single process. On the contrary, it seems likely that a NATO/Pact war would be preceded by months or years of cold war. In such an environment, there would be a long list of preparatory measures that *could* be carried out well before formal mobilization occurred. These could include not only raising the readiness level of low-readiness units, but also such important matters as reevaluating operations plans; resolving issues such as how to share equipment, munitions, and support-related duties; filling out prepositioning sets; acquainting NATO officers with terrain and challenges in areas other than those for which they have long been nominally slated; and perhaps preparing to use units not in the regular force structure for specialized defensive missions.

Another potentially crucial step in the premobilization period might be the opening up of assembly lines to produce munitions and equipment. As noted below, this could be essential. Many observers believe, however, that such steps would be extremely unlikely and that having large common stockpiles is the only sensible approach for NATO to take.

SUSTAINABILITY

Most net assessments do not deal with sustainability issues in depth—rather, they summarize sustainability in terms of nominal days of supply (e.g., NATO's force goals include being able to sustain conventional conflict for n days, where n is always much less than would be prudent, and the assessed actual value of n at a given time is less yet).

Sustainability has many dimensions. First, there is the simple issue of ammunition. It is well known that NATO as a whole has only a limited supply of the high-technology munitions that increase substantially the capability of certain weapon systems. Also, it is well known that certain nations within NATO have shortages of particular munitions. In war games and simulations, these shortages can be treated more or less realistically, although there are differences of opinion about what realism is, since under wartime conditions there would surely be more of an effort to share ammunition than can be exercised in peacetime when NATO is attempting to pressure each of its member nations into fulfilling its obligations.

If conventional defense were successful for a long enough period, replacement supplies of the simpler munitions could be produced, but the initial and potentially decisive phase of war would probably be over before they would become available. A

careful net assessment should account for this explicitly. More study should also be conducted of ways in which high-technology munitions could be produced more quickly than is now possible.

Turning from munitions to other matters, NATO would have trouble replacing both major end items of equipment and the people to man such equipment (e.g., tanks). There are substantial stocks of prepositioned war materiel in Europe, and more stocks exist in the United States and elsewhere, but the history of modern warfare suggests that equipment may suffer high attrition rates, and there is reason to believe it would take quite a long time for the United States and Western Europe to begin producing such equipment—too long to affect results in even a moderately protracted war (see Levin, 1988, for discussion).

Another aspect of sustainability that should be mentioned here is repair. NATO forces have generally emphasized the ability to repair damaged equipment such as tanks relatively far forward; Soviet doctrine deemphasizes this. One tangible measure of the NATO support structure's value is the difference in repair rates assumed for Red and Blue in war games and simulations. In our work at RAND, we do assume such asymmetries, although there are, as always, uncertainties.

In dealing with sustainability, then, we again find that simulation provides outlets for information that should be reflected in net assessments.

IV. CONCLUSIONS

The principal purpose of this Note has been to discuss a new paradigm for conceiving and assessing the military balance and to comment on the role issues of readiness, mobilization, and sustainability can play in such assessments. It is appropriate, however, to spend some time discussing improvement measures, even though no details can be presented here.

A program to improve the Central Region balance should be conceived as a package with components involving force structure, top-level military strategy, operational planning and related political constraints, training, and grand strategy. The balance is multifaceted and should be approached that way.

There are long and well-known lists of possible measures. Without providing details here, I mention a few of these measures, along with some personal comments:[1]

- Increase operational reserves, especially in NATO's weak Northern Army Group (NORTHAG) area. In considering options for doing so, include redeployment of existing forces, development of new units that would be manned by Europeans but equipped by drawing on existing stocks, and the innovative use of light infantry forces to cover specific zones in which infantry is especially effective—thereby releasing mechanized forces for other purposes.

- Expand efforts to provide capability for the creation of obstacles in crisis or wartime—not just on the intra-German border in the form of permanent barriers, but also (and perhaps more importantly) wherever they are needed in the course of combat. That is, consider obstacle creation to be a force multiplier at the tactical and operational level rather than merely a strategic option that is unlikely to be adopted, for political reasons.

- Develop stockpiles adequate for the initial and probably decisive phase of combat. Develop production capability to sustain conventional conflict thereafter as long as necessary to thwart a Pact invasion. This requirement should not be compromised by fuzzily constructed assumptions about the role of nuclear weapons, even though the nuclear deterrent should be preserved and will probably continue to play a dominant role. The quantitative "requirements" for sustainability should be reexamined critically and should reflect uncertainties about the actual intensity of warfare, the distribution of intensity across corps sectors and time, and the feasibility and reliability of intratheater distribution during war. The conclusions should inform judgments about both stockpile requirements and the value of certain support forces.

---

[1]See also, for example, the ESECs report (European Security, 1983), Mearsheimer (1984, 1988, and earlier), Huber (1986), Von Mellenthin and Stolfi (1984), and many other studies.

- Move toward a conception of military strategy that recognizes the *necessity* of having fundamentally different operational strategies for different scenarios. Develop detailed plans for different scenarios and test them in realistic theater-level exercises—even if doing so raises political tensions.

- Modify the interpretation of NATO'S MC 14/3 to require the national partners to develop stockpiles, production capability, and wartime distribution systems adequate to sustain NATO forces in a protracted conventional war. However, be realistic in the assessment of likely warning times.

- Finally, look upon arms control as an opportunity rather than a complication. More than any other factor, arms control has the potential for effecting major changes in the real and perceived balance—for good or for bad.

- In approaching arms control, seek restrictions on Soviet ability to increase the readiness of reserve forces; also, seek large and highly asymmetric reductions (Thomson and Gantz, 1987) and, perhaps, a slow shift toward so-called "defensive defense systems" (e.g., von Bulow in Pierre, 1986, and Huber, 1986).

Appendix A
SELECTED COMMENTS ABOUT ANALYTIC METHODOLOGY

In recent years, a number of attempts have been made to go beyond simple theater-force-ratio analyses and toward more dynamic treatments of warfare, while still using very simple and highly aggregated models. There is always an important place for simplified models in the communication of specific concepts, but serious problems arise when simplified models used to draw broad conclusions omit many of the important factors in the phenomenon being studied. So it is, in my opinion, with the analyses focused on Lanchester equations (e.g., Kaufmann, 1983), the FEBA-expansion model (Posen, 1985, 1988), and certain improvements over Lanchester equations (Epstein, 1985, 1987). Although they can be useful for some purposes, they are not a good basis for assessing the military balance—especially if one believes that maneuver phenomena are important.

GENERAL COMMENTS

There are several ways to study possible war outcomes and the dynamics leading to those outcomes. These include:

- Solving differential equations (e.g., Lanchester equations) analytically—i.e., obtaining closed-form solutions for variables such as the remaining force levels of attacker and defender as a function of time.
- Using computer simulations to describe battle and campaign dynamics.
- Conducting human war games (usually, with computer simulations providing support).

In my view, one must get to the last two methods *to address many of the most important current issues.* The first method depends on assuming a single continuous battle fought to a conclusion under constant conditions.[1] Real warfare, however, even within a corps sector, has the character of a *sequence* of battles under very different circumstances of terrain, defender preparations, defender objectives, air power, natural barriers, and so on. Furthermore, the forces involved in such battles arrive and are pulled

[1]To illustrate this, consider Lanchester's square-law equations, which describe the force strength of attacker and defender over time in terms of the initial strengths and the attacker's and defender's killing-rate coefficients, which are assumed to be constant.

out in discrete chunks, often after losing cohesion rather than being annihilated. Although in abstract theory it might happen that these complications "average out," the essence of operational art is to assure that they do not. Both sides attempt to modulate the intensity of battle so as to maximize intensity in the areas and at the times when they have favorable conditions.[2]

To be sure, individual battles involve attrition, and the laws of combat describing that attrition may be reasonably approximated by simple equations such as Lanchester's or some variant, under many circumstances. Computer simulations use such equations for one time period at a time, but they allow for the coefficients of those equations to change from one time period to the next—thereby reflecting the changes in type of battle, terrain, and so on.

In theory, it might happen that computer simulations would agree with the solutions of differential equations (using appropriate average values of the attacker's and defender's killing coefficients over time). However, except in special cases (what might be termed static attrition warfare), campaign outcomes are more sensitive to the distribution of battle types over time, which is a result of all the operational considerations mentioned above, than to the details of the attrition equations used or the precise way in which forces and their effectiveness are counted and scored. To put it differently, Lanchester-like equations can be very useful for understanding certain types of *local* phenomena, but the character of the whole is not the character of the average local situation (except in very special instances). Efforts to find more and more elegant or rigorous ways to solve equations that do not describe the phenomena at issue represent misdirected effort.

Some of the models used in balance assessments are, in a sense, extremely simple simulations—so simple that they can be applied on the back of the proverbial envelope, or at least with a simple spreadsheet program and a personal computer. The FEBA-expansion model, for example, adjusts the number of forces on line time period by time period to reflect a postulated expansion in the length of the front line over time. Epstein's model allows the forces to adjust their tactics somewhat from one time period to another. Each of these models, however, considers some changes in the nature of the battle over time, but not others. More elaborate simulation models attempt to reflect as many of the factors as seem to be important to the analysis at hand. They quickly become

[2]Epstein (1985) identifies real and serious problems with Lanchester equations, but does not go nearly far enough in correcting them. Instead, he takes only one step down the path toward simulation.

complicated enough that computerization is virtually essential. The simulation models may still be easy to understand and explain, but the computerization greatly simplifies the arithmetic and bookkeeping. (The simulations may, on the other hand, become entirely opaque, i.e., not understandable.

As one more general observation, anecdotes about real wars and real battles often emphasize the role of special circumstances such as a particularly effective general, a successful surprise operation, or the superiority of one weapon system over another given the tactics used. In models and simulations, such special circumstances should be (but seldom are) considered as partially correlated random events. That is, the battles of a war do not all follow the same laws of war with the same parameter values. Instead, defenses sometimes hold and sometimes fail under "the same conditions," but a lucky success now may increase the likelihood of a subsequent success.

As a result, efforts to assess the military balance based on equations (or simulations) that do not account for uncertainties and "random factors" can be misleading and, in my experience, can often be defense-optimistic—especially if the defender is attempting to defend with a marginal force-to-space ratio and minimal reserves, or if the defender is slow to maneuver ground and air forces in response to events. The attacker has advantages in all of this by having the initiative. This is why NATO generals talking about their defense strategy increasingly emphasize the necessity (not merely the desirability) of regaining the initiative as quickly as possible. It is also why Israeli military figures are passionate about the necessity of going onto the offensive as soon as possible. While the famous rules of thumb about defender advantage up to 3:1 and the imagery of Lanchester-like equations suggest that the defender can be static, in fact, it appears that a defender must be very active merely to achieve the effectiveness usually ascribed to him.

A FEW SPECIFIC COMMENTS

Let me next comment on some of the specific problems I see in several published models. These are merely examples, and this is by no means an attempt to be comprehensive. I should also emphasize that *all* models can be criticized (including, most definitely, the simulation models used by my colleagues and me at RAND), and that simple models with shortcomings can nonetheless be quite useful for specific purposes when manipulated by careful analysts. Thus, my comments are not intended to be standard nitpicking, but rather to illustrate how the models in question convey a *picture*

of warfare and the major factors in the military balance that is very different from my own. Briefly, then:

- Models that purport to provide probabilities of defender success (e.g., Kaufmann, 1983) are doing a great disservice, since most of the causes of uncertainty are not reflected in the underlying calculations.[3]

- Models assuming continuous and cohesive defensive lines, even lines that are allowed to stretch, miss important aspects of maneuver phenomena and do not really represent "breakthrough" well. It is notable that the attacker has historically broken approximately even in successful campaigns, rather than suffering an exchange ratio of 2:1 or so as is often assumed in simple models. Often, a key to success was changing the character of the battle from assault to exploitation of breakthrough.

- Models that allow forces to fight to completion, as in closed-form solutions of Lanchester or similar equations, the FEBA-expansion model, and Epstein's model, ignore cohesion-related "breakpoint" phenomena, which have an important effect on outcomes, as can be demonstrated with simulation and as is recognized by military doctrine through echeloning methods and other plans for troop rotation.

- Treating airpower as merely another source of firepower to be added to the combat potential of ground forces largely ignores the counter-maneuver aspects of air operations, which historically have been far more important than air-to-ground attrition (e.g., see Dews and Kozaczka, 1981). Also, it tends to focus attention on close air support, which is probably quite misleading.

I would hasten to note that the authors are aware of most of the above shortcomings and often mention them in footnotes to their articles. Moreover, these shortcomings are by no means unique to the models cited. Again, my purpose is to argue that the simpler models are really missing much of the essence of warfare and constitute a poor basis on which to proceed. While any specific problem can be worked around or repaired within the framework of the simple models, it seems evident that one should either stick with the high-level view (e.g., buildup curves) or move toward real simulation models in which these and many other complications of warfare can be discussed and treated straightforwardly (albeit, with enormous uncertainties, approximations, and, often,

[3]See also articles by Lepingwell (1987) and Homer-Dixon (1987). Both are interesting and informative for those interested in Lanchester equations. They are extremely critical of Kaufmann's work—indeed, unreasonably so in my view, since in some cases they assess analytic sin primarily by noting lack of rigor rather than examining whether heuristic methods (e.g., aggregated measures of combat potential) are convenient and approximately right in the context of real calculations.

errors). Of course, as with most medicines, there can be bad side effects. If the simulations become too complex to understand or if users come to treat them as black-box answer machines, they serve the purposes of analysis poorly. It is a good rule to report no conclusions of complex models unless they can be thoroughly understood in relatively simple terms—at least in retrospect.

Appendix B
MODELS FOR THEATER-LEVEL ANALYTIC WAR GAMING

This appendix provides some information on the models used in the RAND Strategy Assessment Center for Central Region studies.[1] Some details and most of the data pertaining to these models are classified. The model described below has been complemented with human gaming at a higher level of resolution (e.g., IDAHEX) and considerable offline analysis, since it is still rather aggregated in order for it to be useful in policy analysis. As should be evident, the model provides numerous outlets for representing issues of readiness, mobilization, and sustainability.

ENTITIES AND LEVEL OF RESOLUTION

The model follows Blue brigades and Red divisions, by name if necessary. A given unit is *principally* characterized by its score in Equivalent Divisions and Effective Equivalent Divisions (EDs and EEDs)—the former measuring weapon capabilities and the latter including effects of incomplete mobilization, incomplete training, loss of cohesion following intense combat, inefficiencies due to operations in a corps sector where the principal language is different from its own, combat inefficiencies due to using another nation's supplies, and, sometimes, subjective factors for national fighting quality (e.g., if we are attempting to understand a battle in an Arab-Israeli war or a World War II battle on the Eastern Front, it is essential to include quality factors).

The model also keeps track of more detailed information on force composition (e.g., number of tanks, number of artillery tubes, etc.), which it uses in rules designed to avoid some of the more egregious errors of aggregated-firepower methods. For example, a division consisting only of artillery and air defense assets cannot attack. In recent work, we have begun to use situationally adjusted scores, so that, for example, infantry units are assessed as unusually capable in mountains and urban terrain.

Fixed-wing aircraft and helicopters are treated separately,[2] flying a variety of missions with mission- and type-battle-dependent effectiveness and vulnerabilities. Aircraft kill enemy forces *and* slow or disrupt their maneuver. For example, they can

[1]See also, Bennett, Jones, Bullock, and Davis, 1988.

[2]In some analyses, it is also necessary to disaggregate certain other weapon systems of ground forces. The Multiple Launch Rocket System (MLRS), for example, is sometimes treated separately depending on the analysis being conducted and the year for which it applies.

slow the movement rate of a force that has just achieved a breakthrough and could otherwise move at high speeds. They can also reduce the effective force ratio on the forward line of own troops (FLOT) through battlefield interdiction (BAI) missions disrupting the opponent's tactical maneuver, and through both BAI and AI missions delaying the arrival "on the FLOT" of forces being sent to the front from corps or theater-level reserves. The aircraft suffer attrition of various types (on the ground, en route to their mission, during their mission and egress); air-to-ground aircraft also suffer "virtual attrition" in the sense that their effectiveness can drop precipitously if physical attrition is high and conditions are therefore of the type in which historically effectiveness has been very low, despite physical capabilities. Bases can be attacked, reducing sortie rates and destroying unsheltered aircraft.

The "game board" consists of axes of advance broken down into zones of constant character. A zone might be, for example, 40 km on a side and characterized by mixed terrain and other descriptors. Although the model uses axes of advance, forces can maneuver from one axis to another, must protect flanks, and can participate in large-scale envelopment operations involving more than one axis.

In summary, the model's resolution is high in some respects (unit characteristics, type battle, type mission, and so on), but low in others (e.g., with attrition and movement adjudicated for an entire corps in one calculation, and with only some types of support such as tank repair modeled at all).

MAJOR PROCESSES

Forces engage in processes such as alert, deployment to theater, maneuver within the theater, attack, withdrawal, and delay. Engineer assets can produce barriers. Ground forces are subject to attrition from a variety of sources (ground forces on the FLOT, ground forces in the rear, air forces, missiles, etc.). FLOT movement is another major process.

Attrition. Ground-combat attrition for the opposing FLOT forces in a given corps sector is a function of the force strengths (in EEDs), the defender's density (measuring force-to-space ratio), terrain, defender preparations, and type battle. The concept of type battle is especially important, and the types recognized are: static engagement, meeting engagement, assault on hasty defense, assault on deliberate defense, assault on prepared defense, assault on fortified defense, delay, withdrawal, and breakthrough/pursuit. A defender ordered to hold ground with inadequate forces will eventually suffer a breakthrough, after which the type battle is breakthrough/pursuit. The defender's losses

are then very high and the exchange ratio is favorable to the attacker. These types of battle, which have been observed frequently in warfare, are non-Lanchesterian and are often typified by local envelopments with large-scale surrenders or annihilations. Roughly speaking, the attacker suffers grievous casualties during an assault phase but, if he achieves a breakthrough, subsequently gets the opportunity to recoup his losses and more.

The equations used for ground-combat attrition are based on a combination of judgment, historical insights, and analytic convenience. They (as well as parameters of the equations) can be and are varied in sensitivity studies. Overall levels of attrition are lower than in many models and compare favorably with results seen in the Arab-Israeli wars and World War II, including results reported recently (Stoeckli, 1985) regarding Soviet losses on the Eastern Front. Nonetheless, intensity of war (loss rates) is uncertain by as much as a factor of three.

Movement Rates. Movement rates of the FLOT depend on adjusted force ratio, defender density, phase of battle, terrain, the presence of natural or artificial barriers, and the weight and character of airpower. In a variant method that integrates other effects, movement rates depend strongly on the ratio of the sides' loss rates. In either method, typical results are that movement is very slow as long as the defender has a good force-to-space ratio after accounting for attrition and associated loss of cohesion. If the defender density drops, however, movement picks up. A defender who insists on trying to hold ground with inadequate forces will suffer a breakthrough, after which movement is very rapid. The result, then, is that average movement rates over the duration of an army operation can be considerable. The movement rates compare favorably with those observed in World War II, including those from the Eastern Front on which the Soviets were able to emphasize large-scale maneuver. It is by no means clear whether modern movement rates would be faster, slower, or comparable, but the increased mechanization since World War II would not be the critical factor.

Movement. Ground forces are maneuvered within and across sectors, both in accordance with initial plans and as the result of subsequent adaptations. With some human interaction to guard against model problems, it is possible to simulate large-scale envelopments such as those emphasized in Soviet doctrine and demonstrated repeatedly in their World War II experience.

Within a corps sector, forces are sent to the front, pulled out for recovery when their cohesion has dropped due to attrition (unless pulling them out would leave the line uncovered), and moved back up as appropriate. Unit cohesion drops faster than attrition

occurs to reflect "breakpoint" phenomena, which are often very important to results. Command-control decisions can be fully or partially automated, although it is always desirable to have an analyst reviewing the model's decisions on a day-by-day basis. Command-control delays and imperfections are modeled, based in part on experience from human war gaming in which such effects can be crucial (and in which a side often does much worse than its force levels would suggest it "should" from a force-ratio analysis).

ANALTYIC WAR PLANS

An especially useful feature of our approach is that it includes expicit Red and Blue "analytic war plans." These models govern operations of ground and air forces consistent with the sides' theater-level strategies *and* the need to adapt orders to circumstances as the campaign unfolds. Higher-level aspects of the strategies are inputs (e.g., objectives, priorities, and the basic scheme of maneuver) that can be adjusted interactively by human players or the analyst, but the translation of strategy into specific force orders such as where to commit operational reserves over time, or how to adjust the focus of tactical air operations, is accomplished through heuristic rules based on experience in previous human games, plus subsequent analysis that attempts to cover more of the "what if?" situations than have actually been observed in games. Over time, we are developing a library of such analytic war plans to represent a range of different operational strategies. These plans are often developed in one study, and then used in a number of subsequent studies.

BIBLIOGRAPHY

Almquist, Peter (1987), "Moscow's Conventional Wisdom: Soviet Views of the European Balance," *Arms Control Today*, December, p. 20 (cited in Levin 1988).

Battilega, John A., and Judith K. Grange (eds.) (1984), *The Military Applications of Modeling*, Air Force Institute of Technology Press, Wright-Patterson Air Force Base, Ohio.

Bennett, Bruce W., Carl M. Jones, Arthur Bullock, and Paul Davis (1988), *Main Theater Warfare Modeling in the RAND Strategy Assessment System (3.0)*, The RAND Corporation, N-2743-NA, September.

Betts, Richard K. (1982), *Surprise Attack*, Brookings Institution, Washington, D.C.

Biddle, Stephen D. (1988), "The European Conventional Balance," *Survival*, March/April.

Blaker, James, and Andrew Hamilton (1977), "Assessing Military Balances: The NATO Example," reprinted in John Reichart and Steven Sturm (eds.), *American Defense Policy*, 5th ed., The Johns Hopkins University Press, Baltimore, Md

Caidin, Martin (1974), *The Tigers Are Burning*, Hawthorn Books, New York.

Canby, Steven (1986), "NATO Defense: What Can Be Done," unpublished; see also *The Alliance and Europe: Part IV, Military Doctrine and Technology*, Adelphi Papers 109, The International Institute for Strategic Studies, London, UK.

Davis, Paul K. (1982), *Observations on the Rapid Deployment Joint Task Force: Origins, Direction, and Mission*, The RAND Corporation, P-6751.

Davis, Paul K. (1987), *Analytic War Gaming with the RAND Strategy Assessment System (RSAS)*, The RAND Corporation, Research Brief-7801.

Davis, Paul K. (1985), *Game-Structured Analysis as a Framework for Defense Planning*, The RAND Corporation, P-7051; published originally as a chapter in Reiner Huber (ed.) (1986), *Modeling and Analysis of Conventional Defense in Europe*, Plenum Publishing Corporation, New York.

Davis, Paul K. (forthcoming), *Toward a Conceptual Framework for Operational Arms Control in Europe's Central Region*, The RAND Corporation, R-2704-USDP.

Department of Defense (1987), *Soviet Military Power*, Government Printing Office, Washington, D.C.

Dews, Edmund, and Felix Kozaczka (1981), *Air Interdiction: Lessons from Past Campaigns*, The RAND Corporation, N-1743-PA&E.

Donnelly, C. N. (1983), *The Soviet Threat to Europe in the 1980's*, Soviet Studies Research Centre, Sandhurst, UK.

Dupuy, T. N. (1979), *Numbers, Predictions, and War*, Bobbs-Merrill, Indianapolis.

Enthoven, Alain C., and K. W. Smith (1971), *How Much Is Enough? Shaping the Defense Program, 1961-1969*, Harper and Row, New York.

Epstein, Joshua (1985), *The Calculus of Conventional War*, Brookings Institution, Washington, D.C.

Epstein, Joshua (1987), *The 1988 Defense Budget*, Brookings Institution, Washington, D.C.

European Security Study (1983), *Strengthening Conventional Defense in Europe*, St. Martin's Press, New York.

Homer-Dixon, Thomas F. (1987), "A Common Misapplication of the Lanchester Square Law: A Research Note," *International Security*, Summer.

Huber, Reiner (ed.) (1986), *Modeling and Analysis of Conventional Defense in Europe*, Plenum Publishing Corporation, New York.

Karber, Phillip A. (1984), "In Defense of Forward Defense," *Armed Forces Journal International*, Washington, D.C.

Kaufmann, William W. (1983), "The Arithmetic of Force Planning," in John D. Steinbruner and Leonard V. Sigal (eds.), *Alliance Security: NATO and the First Use Question*, Brookings Institution, Washington, D.C.

Lepingwell, John W. R. (1987), "The Laws of Combat: Lanchester Reexamined," *International Security*, Summer.

Levin, Carl (1988), *Beyond the Bean Count: Realistically Assessing the Conventional Military Balance in Europe*, Report to the U.S. Senate Committee on Armed Services, January 19.

Mako, William (1983), *U.S. Ground Forces and the Defense of Central Europe*, Brookings Institution, Washington, D.C.

Mearsheimer, John J. (1984), "Numbers, Strategy, and the European Balance," *International Security*, Spring (Vol. 12, No. 4).

Mearsheimer, John J. (1988), *International Security*, Spring (Vol. 12, No. 4).

Pierre, Andrew J. (ed.) (1986), *The Conventional Defense of Europe: New Technologies and New Strategies*, Council on Foreign Relations, New York.

Posen, Barry (1985), "Competing Views of the Central Region Conventional Balance," in Keith Dunn and William O. Standenmaier (eds.), *Military Strategy in Transition: Defense and Deterrence in the 1980s*, Westview Press, Inc., New York.

Posen, Barry (1988), "Is NATO Decisively Outnumbered?" *International Security* (Vol. 12, No. 4).

Shilling, David (1988), "Solid Progress, But Challenges Remain," *Survival*, March/April.

Stoeckli, Major H. G. (1985), *Soviet Operational Planning: Superiority Ratios vs. Casualty Rates*, Soviet Studies Research Center, Sandhurst, UK.

Thomson, James A., and Nanette C. Gantz (1987), *Conventional Arms Control Revisited: Objectives in the New Phase*, The RAND Corporation, N-2697-AF.

Vigor, Peter H. (1983), *Soviet Blitzkrieg Theory*, St. Martin's Press, New York.

Von Mellenthin, F. W. (1955), *Panzer Battles: A Study of the Employment of Armour in the Second World War*, Futura Publications, London, reprinted 1984.

Von Mellenthin, F. W., and R. H. S. Stolfi (1984), *NATO Under Attack: Why the Western Alliance Can Fight Outnumbered and Win in Central Europe Without Nuclear Weapons*, Duke University Press, Durham, N. C.